What is claimed is:

1. A method for constructing a variant set for an antibody of interest, the method
5   comprising:

a) identifying, using a plurality of rules, a plurality of positions in said
antibody of interest and, for each respective position in said plurality of positions, one
or more substitutions for the respective position, wherein the plurality of positions and
the one or more substitutions for each respective position in the plurality of positions
10   collectively define an antibody sequence space;

b) selecting a variant set, wherein said variant set comprises a plurality of
variants of said antibody of interest and wherein said variant set is a subset of said
antibody sequence space;

c) measuring a property of all or a portion of the variants in said variant set;
15   d) modeling a sequence-activity relationship between (i) one or more
substitutions at one or more positions of the antibody of interest represented by the
variant set and (ii) the property measured for all or said portion of the variants in the
variant set; and

e) redefining said variant set to comprise variants that include substitutions in
20   said plurality of positions that are selected based on a function of said sequence-
activity relationship.

2. The method of claim 1, the method further comprising repeating said measuring,
modeling, and, optionally, said redefining, until a variant in said variant set exhibits a
25   value for said property that exceeds a predetermined value.

3. The method of claim 2 wherein said predetermined value is a value that is greater
than the value for the property that is exhibited by said antibody of interest.

30   4. The method of claim 1, the method further comprising repeating said measuring,
modeling, and, optionally, said redefining, until a variant in said variant set exhibits a
value for said property that is less than a predetermined value.

5. The method of claim 4 wherein said predetermined value is a value that is less than the value for the property that is exhibited by said antibody of interest.

6. The method of claim 1, the method further comprising repeating said measuring,
5   modeling, and, optionally, said redefining, a predetermined number of times.

7. The method of claim 6 wherein said predetermined number of times is two, three, four, or five.

10   8. The method of claim 1 wherein said sequence-activity relationship comprises a plurality of values and wherein each value in said plurality of values describes a relationship between (i) a substitution at a position in said plurality of positions represented by said all or said portion of the variants in said variant set and said property, (ii) a plurality of substitutions at a position in said plurality of positions
15   represented by said all or said portion of the variants in said variant set and said property, or (iii) one or more substitutions in one or more positions in said plurality of positions represented by said all or said portion of the variants in said variant set and said property.

20   9. The method of claim 8 wherein said modeling comprises regressing:

$$V_{measured} = W_{11}P_1S_1 + W_{12}P_1S_2 + \quad + W_{1N}P_1S_N + \quad + W_{M1}P_MS_1 + W_{M2}P_MS_2 + \quad + W_{MN}P_MS_N$$

wherein,

25   $V_{measured}$ represents the property measured in variants in said variant set;

W_{MN} = is a value in said plurality of values;

P_M = is a position in said antibody of interest in said plurality of positions in said antibody of interest; and

30   S_N = is a substitution in the one or more positions for a position in the plurality of positions in said antibody of interest.

10. The method of claim 9 wherein said regressing comprises linear regression, non-linear regression, logistic regression, multivariate data analysis, or partial least squares projection to latent variables.

5   11. The method of claim 1 wherein said modeling comprises computation of a neural network, computation of a bayesian model, a generalized additive model, a support vector machine, or classification using a regression tree.

12. The method of claim 1 wherein said modeling comprises boosting or adaptive

10   boosting.

13. The method of claim 1 wherein said redefining further comprises:

computing a predicted score for a population of variants of said antibody of interest using said sequence-activity relationship, wherein each variant in said

15   population of variants includes a substitution at one or more positions in said plurality of positions in said antibody of interest; and

selecting said variant set from among said population of variants as a function of the predicted score received by each variant in said set of variants.

20   14. The method of claim 13, the method further comprising

ranking said population of variants, wherein each variant in said population of variants is ranked based on the predicted score received by the variant based upon the sequence-activity relationship; and

said selecting comprising accepting a predetermined percentage of the top

25   ranked variants in said population of variants for said variant set.

15. The method of claim 13, wherein a respective variant in said population of variants is selected for said variant set when the predicted score of the respective variant exceeds a predetermined value.

30

16. The method of claim 1 wherein said redefining step (e) further comprises redefining said variant set to comprise one or more variants each having a substitution in a position in said plurality of positions not present in any variant in the variant set selected by said selecting step (b).

17. The method of claim 1 wherein

said modeling a sequence-activity relationship (d) further comprises modeling a plurality of sequence-activity relationships, wherein each respective sequence-activity relationship in said plurality of sequence-activity relationships describes the relationship between (i) one or more substitutions at one or more positions of the antibody of interest represented by the variant set and (ii) the property measured for all or said portion of the variants in the variant set; and

said redefining said variant set (e) comprises redefining said variant set to comprise variants that include substitutions in said plurality of positions that are selected based on a combination of said plurality of sequence-activity relationships.

18. The method of claim 17, the method further comprising:

repeating said measuring based upon said redefined variant set, wherein a property of all or a portion of the variants in the redefined variant set is measured; and

weighting each respective sequence-activity relationship in said plurality of sequence activity relationships based on an agreement between (i) measured values for the property of variants in said redefined variant set and (ii) values for the property of variants in said redefined variant set that were predicted by said respective sequence-activity relationship, wherein

a first sequence-activity relationship that achieves better agreement between measured and predicted values than a second sequence-activity relationship receives a higher weight than said second sequence-activity relationship.

19. The method of claim 17 wherein said redefining step (e) further comprises redefining said variant set to comprise one or more variants each having a substitution in a position in said plurality of positions not present in any variant in the variant set selected by said selecting step (b).

20. The method of claim 18 wherein said redefining step (e) further comprises redefining said variant set to comprise one or more variants each having a substitution in a position in said plurality of positions not present in any variant in the variant set selected by said selecting step (b).

21. The method of claim 1, wherein

the contribution of each respective rule in said plurality of rules to said antibody sequence space is independently weighted by a rule weight in a plurality of rule weights corresponding to the respective rule; and

5   the method further comprising, prior to said redefining step (e), the steps of:

adjusting one or more rule weights in said plurality of rule weights based on a comparison, for each respective variant in the variant set, (i) a value assigned to the respective variant by said sequence-activity relationship, and (ii) a score assigned by the plurality of rules to the respective variant; and

10   repeating said identifying step using said rule weights, thereby redefining said plurality of positions and, for each respective position in said plurality of positions, redefining the one or more substitutions for the respective position; and wherein

said redefining step (e) further comprises redefining said variant set to

15 comprise one or more variants each having a substitution in a position in said redefined plurality of positions not present in any variant in the variant set selected by said initial selecting step (b).

22. The method of claim 21 wherein

20   · said modeling a sequence-activity relationship (d) further comprises modeling a plurality of sequence-activity relationships, wherein each respective sequence-activity relationship in said plurality of sequence-activity relationships describes the relationship between (i) one or more substitutions at one or more positions of the antibody of interest represented by the variant set and (ii) the property measured for

25 all or said portion of the variants in the variant set; and

said redefining said variant set (e) comprises redefining said variant set to comprise variants that include substitutions in said plurality of positions that are selected based on a combination function of said plurality of sequence-activity relationships.

30

23. The method of claim 22, the method further comprising:

repeating said measuring based upon said redefined variant set, wherein a property of all or a portion of the variants in the redefined variant set is measured; and

weighting each respective sequence-activity relationship in said plurality of sequence activity relationships based on an agreement between (i) measured values for the property of variants in said redefined variant set and (ii) values for the property of variants in said redefined variant set that were predicted by said respective

5   sequence-activity relationship, wherein

a first sequence-activity relationship that achieves better agreement between measured and predicted values than a second sequence-activity relationship receives a higher weight than said second sequence-activity relationship.

10   24. The method of claim 1 wherein said antibody of interest is IgG, IgE, IgM, IgD, IgA or IgY.

25. The method of claim 1 wherein said plurality of positions comprises five or more positions.

15

26. The method of claim 1 wherein said plurality of positions comprises ten or more positions.

27. The method of claim 1 wherein said plurality of rules comprises two or more

20   rules.

28. The method of claim 1 wherein said plurality of rules comprises five or more rules.

25   29. The method of claim 1 wherein, a rule in said plurality of rules assigns a score to a variant of said antibody of interest by considering a lineup of a plurality of sequences that are homologous to said antibody of interest.

30. The method of claim 1 wherein, a rule in said plurality of rules assigns a score to

30   a variant of said antibody of interest by considering structural variations in one or more three dimensional structures of antibodies that are homologous to said antibody of interest.

31. The method of claim 1 wherein a rule in said plurality of rules assigns a score to a variant of said antibody of interest by considering variations in a substitution matrix for said antibody of interest.

5   32. A population of cells comprising nucleotide sequences encoding variants in the redefined variant set of an instance of step e) of claim 6.

33. A population of cells comprising variants in the redefined variant set of an instance of step e) of claim 6.

10

34. The method of claim 1, wherein said identifying combines a score from each rule in said plurality of rules for a variant of an antibody of interest.

35. The method of claim 34 wherein said combining comprises adding (i) a first score

15   from a first rule in said plurality rules and (ii) a second score from a second rule in said plurality rules for said variant of an antibody of interest.

36. The method of claim 34 wherein said combining comprises multiplying (i) a first score from a first rule in said plurality rules and (ii) a second score from a second rule

20   in said plurality rules for said variant of an antibody of interest.

37. The method of claim 1 wherein said variant set consists of between 5 and 200 variants of said antibody of interest.

25   38. The method of claim 1 wherein said variant set consists of between 15 and 50 variants of said antibody of interest.

39. The method of claim 1 wherein said selecting said variant set (b) comprises applying a monte carlo algorithm, a genetic algorithm, or a combination thereof, to

30   construct said variant set, with the provisos that:

(i) each variant in all or portion of said variant set has a number of substitutions that is between a first value and a second value; and

(ii) a number of different pairs of substitutions collectively represented by said variant set is above a predetermined number.

40. The method of claim 39 wherein said first value is two substitutions and said second value is twenty substitutions.

5   41. The method of claim 39 wherein said first value is four substitutions and said second value is ten substitutions.

42. The method of claim 39 wherein said predetermined number is thirty.

10   43. The method of claim 39 wherein said predetermined number is one hundred.

44. The method of claim 1 wherein said selecting said variant set (b) comprises:

dividing said antibody of interest into one or more functional domains; and for each respective functional domain in said one or more functional domains,

15   applying a monte carlo algorithm, a genetic algorithm, or a combination thereof, in order to identify substitutions at positions in the plurality of positions that are in said respective functional domain for inclusion in one or more variants in said variant set, with the provisos that:

all or a portion of the variants in the variant set contains a predetermined

20   number of substitutions at positions from each of the one or more functional domains; and

a number of different pairs of substitutions at positions in each of the one or more functional domains that is collectively represented by the variant set is above a threshold value.

25

45. The method of claim 44 wherein said predetermined number is between two and twenty.

46. The method of claim 44 wherein said predetermined number is between four and

30   ten.

47. The method of claim 44 wherein said threshold value is thirty.

48. The method of claim 44 wherein said threshold value is one hundred.

49. The method of claim 1 wherein

said measuring comprises synthesizing all or said portion of the variants in said variant set, and wherein

5   said property of a variant in said variant set is a level of expression of said variant in a host cell, a susceptibility of said variant to a post-translational modification, a killing of a pathogenic organism or a virus resulting from an activity of said variant, a modulation of a signaling pathway by said variant, a modulation of surface density of a cell-surface receptor by said variant, a binding of a cellular

10   growth factor receptor by said variant, a binding of a receptor or a mediator of tumor-driven angiogenesis by said variant, a binding of a B cell surface antigen by said variant, a binding of a protein synthesized by said variant, an induction of an antibody-mediated cell killing by said variant, an induction of an antibody-dependent macrophage activity by said variant, an induction of a histamine release by said

15   variant, an induction of or cross-reaction with an anti-idiotype antibody by said variant, an immunogenicity of said variant, a reduction of viral titer by said variant or an immunomodulatory activity of said variant.

50. The method of claim 1 wherein said sequence-activity relationship has the form:

20                                  $Y = f(w_1x_1, w_2x_2, \quad w_ix_i)$

wherein,

Y is a quantitative measure of said property;

x_i is a descriptor of a substitution, a combination of substitutions, or a component of one or more substitutions, at one or more positions in said plurality of

25   positions;

w_i is a weight applied to descriptor x_i; and

f( ) is a mathematical function.

51. The method of claim 50 wherein said modeling comprises regressing:

30                                  $Y = f(w_1x_1, w_2x_2, \quad w_ix_i)$.

52. The method of claim 51 wherein regressing comprises linear regression, non-linear regression, logistic regressing, or partial least squares projection to latent variables.

53. The variant of claim 2 that exhibits a value for said property that exceeds a predetermined value.

5   54. The variant of claim 53 wherein said predetermined value is a value that is greater than the value for the property that is exhibited by said antibody of interest.

55. A composition comprising the variant of claim 44, and a carrier.

10   56. The variant of claim 2 that exhibits a value for said property that is less than a predetermined value.

57. The variant of claim 56 wherein said predetermined value is a value that is less than the value for the property that is exhibited by said antibody of interest.

15

58. A composition comprising the variant of claim 56, and a carrier.

59. A plurality of nucleic acid sequences comprising nucleotide sequences encoding all or a portion of the variants in the redefined variant set of step e) of claim 1.

20

60. All or a portion of the variants in the redefined variant set of step e) of claim 1.

61. A plurality of nucleic acid sequences comprising nucleotide sequences encoding variants in the redefined variant set of an instance of step e) of claim 2.

25

62. All or a portion of the variants in the redefined variant set of an instance of step e) of claim 2.

63. A plurality of nucleic acid sequences comprising nucleotide sequences encoding 30   variants in the redefined variant set of an instance of step e) of claim 4.

64. The variants in the redefined variant set of an instance of step e) of claim 4.

65. A plurality of nucleic acid sequences comprising nucleotide sequences encoding variants in the redefined variant set of an instance of step e) of claim 6.

66. The variants in the redefined variant set of an instance of step e) of claim 6.

67. A population of cells comprising nucleic acid sequences encoding a plurality of variants in the redefined variant set of step e) of claim 1.

68. A population of cells comprising the variants in the redefined variant set of step e) of claim 1.

69. A population of cells comprising nucleic acid sequences encoding variants in the redefined variant set of an instance of step e) of claim 2.

70. A population of cells comprising variants in the redefined variant set of an instance of step e) of claim 2.

71. A population of cells comprising nucleotide sequences encoding variants in the redefined variant set of an instance of step e) of claim 4.

72. A population of cells comprising nucleotide sequences encoding variants in the redefined variant set of an instance of step e) of claim 4.

73. The method of claim 1 wherein said antibody of interest is mammalian.

74. The method of claim 1 wherein said antibody of interest is from rat, mouse, chicken, cow, monkey, pig, dog, or rabbit.

75. The method of claim 1 wherein said antibody of interest is a monoclonal antibody, a bispecific antibody, a multispecific antibody, a humanized antibody, a chimeric antibody, a camelised antibody, a single domain antibody, a single-chain Fvs (ScFv), a single chain antibody, a Fab fragment, a F(ab') fragment, a disulfide-linked Fvs (sdFv), or an anti-idiotypic (anti-Id) antibody.

76. The method of claim 1 wherein said antibody of interest is an epitope-binding fragment of a monoclonal antibody, an epitope-binding fragment of a bispecific antibody, an epitope-binding fragment of a multispecific antibody, an epitope-binding fragment of a humanized antibody, an epitope-binding fragment of a chimeric

5   antibody, an epitope-binding fragment of a camelised antibody, an epitope-binding fragment of a single domain antibody, an epitope-binding fragment of a single-chain Fvs (ScFv), an epitope-binding fragment of a single chain antibody, an epitope-binding fragment of a Fab fragment, an epitope-binding fragment of a F(ab') fragment, an epitope-binding fragment of a disulfide-linke Fvs (sdFv), or an epitope-

10   binding fragment of an anti-idiotypic (anti-Id) antibody.

77. The method of claim 1 wherein said antibody of interest is an antibody fragment.

78. The method of claim 1 wherein a variant in the variant set comprises a

15   monoclonal antibody, a bispecific antibody, a multispecific antibody, a humanized antibody, a chimeric antibody, a camelised antibody, a single domain antibody, a single-chain Fvs (ScFv), a single chain antibody, a Fab fragment, a F(ab') fragment, a disulfide-linked Fvs (sdFv), or an anti-idiotypic (anti-Id) antibody.

20   79. The method of claim 1 wherein a variant in the variant set comprises an epitope-binding fragment of a monoclonal antibody, an epitope-binding fragment of a bispecific antibody, an epitope-binding fragment of a multispecific antibody, an epitope-binding fragment of a humanized antibody, an epitope-binding fragment of a chimeric antibody, an epitope-binding fragment of a camelised antibody, an epitope-

25   binding fragment of a single domain antibody, an epitope-binding fragment of a single-chain Fvs (ScFv), an epitope-binding fragment of a single chain antibody, an epitope-binding fragment of a Fab fragment, an epitope-binding fragment of a F(ab') fragment, an epitope-binding fragment of a disulfide-linke Fvs (sdFv), or an epitope-binding fragment of an anti-idiotypic (anti-Id) antibody.

30

80. The method of claim 1 wherein a variant in said variant set comprises an antibody fragment.

81. The method of claim 1 wherein said measuring said property of all or said portion

134

of the variants in the variant set comprises:

 expressing a variant the variant set in a cell line; and

 measuring a cell-surface receptor surface density of said cell line that includes said variant.

5

82. The method of claim 1 wherein said measuring said property of all or said portion of the variants in the variant set comprises:

 expressing a variant in the variant set in a cell line; and

 measuring a cell surface receptor internalization rate of said cell line that

10 includes said variant.

83. The method of claim 1 wherein said measuring said property of all or said portion of the variants in the variant set comprises:

 expressing a variant the variant set in a cell line; and

15 measuring a cell surface receptor post-translational modification of said cell line that includes said variant.

84. The method of claim 83 wherein said cell surface receptor post-translational modification is phosphorylation.

20

85. The method of claim 1 wherein said measuring said property of all or said portion of the variants in the variant set comprises:

 expressing a variant in said all or said portion of the variant set in a cell line; and

25 measuring a binding of an antigen to said cell line that includes said variant.

86. The method of claim 85 wherein said antigen is a cellular growth factor receptor, a receptor of tumor-driven angiogenesis, a mediator of tumor-driven angiogenesis, a B cell surface antigen, or a protein synthesized by or in response to a pathogen.

30

87. The method of claim 1 wherein said measuring comprises measuring the ability for a variant in said variant set to immunospecifically bind to an antigen.

88. The method of claim 87 wherein said measuring comprises placing said variant in solution, spotting said variant onto a microchip, placing a polynucleotide encoding said variant in bacteria, placing a polynucleotide that codes for said variant in a spore, placing a polynucleotide that codes for said variant in a plasmid, or placing a
5 polynucleotide that codes for said variant in phage.

89. The method of claim 1 wherein said measuring said property comprises assaying for a reduction of a viral titer in infected tissue culture cells by a variant in all or said portion of the variant set.

10

90. the method of claim 89 wherein the virus is hepatitis A virus, hepatitis B virus, hepatitis C virus, human papilloma virus, human immunodeficiency virus, respiratory syncitial virus, human adenovirus, fowl adenovirus 1, African swine fever virus, lymphocytic choriomeningitis virus, ippy virus, lassa virus, equine arteritis virus,
15 human astrovirus 1, autographa californica nucleopolyhedrovirus, plodia interpunctella granulovirus, commelina yellow mottle virus, rice tungro bacilliform virus, mushroom bacilliform virus, infectious pancreatic necrosis virus, infectious bursal disease virus, drosophila x virus, alfalfa mosaic virus, tobacco streak virus, brome mosaic virus, cucumber mosaic virus, apple stem grooving virus, carnation
20 latent virus, cauliflower mosaic virus, chicken anemia virus, beet yellows virus, cowpea mosaic virus, tobacco ringspot virus, avian infectious bronchitis virus, alteromonas phage pm2, pseudomonas phage phi6, hepatitis delta virus, carnation ringspot virus, red clover necrotic mosaic virus, sweet clover necrotic mosaic virus, pea enation mosaic virus, ebola virus zair, soil-borne wheat mosaic virus, beet
25 necrotic yellow vein virus, sulfobolus virus 1, maize streak virus, beet curly top virus, bean golden mosaic virus , duck hepatitis B virus, human herpesvirus, human herpesvirus, ateline herpesvirus 2, barley stripe mosaic virus, cryphonectria hypovirus 1-ep713,raspberry bushy dwarf virus, acholeplasma phage 151, chilo iridescent virus, goldfish virus 1, enterobacteria phage ms2, enterobacteria phage qbeta, thermoproteus
30 virus 1, maize chlorotic mottle virus, maize rayado fino virus, coliphage phix174, spiromicrovirus, spiroplasma phage, bdellomicrovirus, bdellovibrio phage, chlamydiamicrovirus, chlamydia phage 1, coliphage t4, tobacco necrosis virus, nodamura virus, influenzavirus a, influenzavirus C, thogoto virus, rabbit (shope) papillomavirus, human parainfluenza virus, measles virus, rubulavirus, mumps virus,

human respiratory syncytial virus, gaeumannomyces graminis virus, penicillium
chrysogenum virus, white clover cryptic virus, white clover cryptic virus 2, minute
mice virus, adeno-associated virus, junonia coenia densovirus, bombyx mori virus,
aedes aegypti densovirus, 1-paramecium bursaria chlorella nc64a virus, paramecium

5   bursaria chlorella virus, 2-paramecium bursaria chlorella pbi virus, 3-hydra viridis
chlorella virus, human poliovirus 1, human rhinovirus 1A, hepatovirus,
encephalomyocarditis virus, foot-and-mouth disease virus, acholeplasma phage 12,
coliphage t7, campoletis sonorensis virus, cotesia melanoscela virus, potato virus X,
potato virus Y, ryegrass mosaic virus, barley yellow mosaic virus, fowlpox virus,

10   sheep pox virus, swinepox virus, molluscum contagiosum virus, yaba monkey tumor
virus, entomopoxvirus A, melolontha melolontha entomopoxvirus, amsacta moorei
entomopoxvirus, chironomus luridus entomopoxvirus, reovirus 3, epizootic
hemarrhogic disease virus 1, or simian rotavirus SA11.

15   91. The method of claim 1 wherein said measuring said property comprises assaying
for a reduction of a viral titer in an animal model by a variant in all or said portion of
the variant set.

92. The method of claim 91 wherein the virus is is hepatitis A virus, hepatitis B virus,

20   hepatitis C virus, human papilloma virus, human immunodeficiency virus, respiratory
syncitial virus, human adenovirus, fowl adenovirus 1, African swine fever virus,
lymphocytic choriomeningitis virus, ippy virus, lassa virus, equine arteritis virus,
human astrovirus 1, autographa californica nucleopolyhedrovirus, plodia
interpunctella granulovirus, commelina yellow mottle virus, rice tungro bacilliform

25   virus, mushroom bacilliform virus, infectious pancreatic necrosis virus, infectious
bursal disease virus, drosophila x virus, alfalfa mosaic virus, tobacco streak virus,
brome mosaic virus, cucumber mosaic virus, apple stem grooving virus, carnation
latent virus, cauliflower mosaic virus, chicken anemia virus, beet yellows virus,
cowpea mosaic virus, tobacco ringspot virus, avian infectious bronchitis virus,

30   alteromonas phage pm2, pseudomonas phage phi6, hepatitis delta virus, carnation
ringspot virus, red clover necrotic mosaic virus, sweet clover necrotic mosaic virus,
pea enation mosaic virus, ebola virus zair, soil-borne wheat mosaic virus, beet
necrotic yellow vein virus, sulfobolus virus 1, maize streak virus, beet curly top virus,
bean golden mosaic virus , duck hepatitis B virus, human herpesvirus, human

herpesvirus, ateline herpesvirus 2, barley stripe mosaic virus, cryphonectria hypovirus 1-ep713, raspberry bushy dwarf virus, acholeplasma phage 151, chilo iridescent virus, goldfish virus 1, enterobacteria phage ms2, enterobacteria phage qbeta, thermoproteus virus 1, maize chlorotic mottle virus, maize rayado fino virus, coliphage phix174,

5   spiromicrovirus, spiroplasma phage, bdellomicrovirus, bdellovibrio phage, chlamydiamicrovirus, chlamydia phage 1, coliphage t4, tobacco necrosis virus, nodamura virus, influenzavirus a, influenzavirus C, thogoto virus, rabbit (shope) papillomavirus, human parainfluenza virus, measles virus, rubulavirus, mumps virus, human respiratory syncytial virus, gaeumannomyces graminis virus, penicillium

10   chrysogenum virus, white clover cryptic virus, white clover cryptic virus 2, minute mice virus, adeno-associated virus, junonia coenia densovirus, bombyx mori virus, aedes aegypti densovirus, 1-paramecium bursaria chlorella nc64a virus, paramecium bursaria chlorella virus, 2-paramecium bursaria chlorella pbi virus, 3-hydra viridis chlorella virus, human poliovirus 1, human rhinovirus 1A, hepatovirus,

15   encephalomyocarditis virus, foot-and-mouth disease virus, acholeplasma phage 12, coliphage t7, campoletis sonorensis virus, cotesia melanoscela virus, potato virus X, potato virus Y, ryegrass mosaic virus, barley yellow mosaic virus, fowlpox virus, sheep pox virus, swinepox virus, molluscum contagiosum virus, yaba monkey tumor virus, entomopoxvirus A, melolontha melolontha entomopoxvirus, amsacta moorei

20   entomopoxvirus, chironomus luridus entomopoxvirus, reovirus 3, epizootic hemarrhogic disease virus 1, or simian rotavirus SA11.

93. The method of claim 1 wherein said measuring said property comprises assaying for a change in rate of proliferation of cells grown in a culture by a variant in all or

25   said portion of the variant set.

94. The method of 93 wherein the culture cells is tumor cells, a cell line derived from tumor cells, a cell line derived from breast cancer cells, a cell line derived from ovarian cancer cells, a cell line derived from lung cancer cells, a cell line derived from

30   bone cancer cells, a cell line derived from fibroblast cancer cells, a cell line derived from hematopoetic cancer cells, a cell line derived from testicular cancer cells, a cell line derived from colon cancer cells, a cell line derived from prostate cancer cells, or a cell line derived from leukemia cells.

95. The method of claim 1 wherein said measuring said property comprises assaying for a change in rate of proliferation of a specific cell type in an animal model by a variant in all or said portion of the variant set.

5   96. The method of 95 wherein the specific cell type is a tumor cell type.

97. The method of claim 96 wherein the specific cell type is derived from a breast cancer tumor, an ovarian cancer tumor, a lung cancer tumor, a bone cancer tumor, a fibroblast cancer, a hematopoetic cancer, a testicular cancer, a colon cancer, a prostate

10   cancer, or a leukemia.

98. A method of preventing, treating or ameliorating one or more symptoms associated with a condition in a subject that is caused by a pathogenic organism, said method comprising administering a prophylactically or therapeutically effective

15   amount of the variant of claim 2.

99. A method of preventing, treating or ameliorating one or more symptoms associated with a condition in a subject that is caused by a virus, said method comprising administering a prophylactically or therapeutically effective amount of the

20   variant of claim 2.

100. A method of preventing, treating or ameliorating one or more symptoms associated with a condition in a subject that is caused a cancer, said method comprising administering a prophylactically or therapeutically effective amount of the

25   variant of claim 2.

101. A method of weighting a plurality of selection rules for use in selecting a plurality of positions in an antibody of interest and, for each respective position in said plurality of positions, one or more substitutions for the respective positions, the

30   method comprising:

        a) identifying, using said plurality of selection rules, a plurality of positions in an antibody of interest and, for each respective position in said plurality of positions, one or more substitutions for the respective position, wherein

the plurality of positions and the one or more substitutions for each respective position in the plurality of positions collectively define an antibody sequence space, and

the contribution of each respective rule in said plurality of rules to said

5   antibody sequence space is independently weighted by a rule weight in a plurality of rule weights corresponding to the respective rule;

b) selecting a variant set, wherein said variant set comprises a plurality of variants of said antibody of interest and wherein said variant set is a subset of said antibody sequence space;

10   c) measuring a property of all or a portion of the variants in said variant set;

d) modeling a sequence-activity relationship between (i) one or more substitutions at one or more positions of the antibody of interest represented by the variant set and (ii) the property measured for all or said portion of the variants in the variant set; and

15   e) adjusting one or more rule weights in said plurality of rule weights based on a comparison, for each respective variant in the variant set, (i) a value assigned to the respective variant by the sequence-activity relationship, and (ii) a score assigned by the plurality of rules to the respective variant;

f) repeating said identifying, selecting, measuring, modeling, and adjusting for

20   each antibody of interest in a plurality of antibodies of interest.

102. The method of claim 101, the method further comprising, prior to said repeating (f):

(i) modeling a sequence-activity relationship between (a) one or more

25   substitutions at one or more positions of the antibody of interest represented by the variant set and (b) the property measured for all or said portion of the variants in the variants in the variant set; and

(ii) redefining said variant set to comprise variants that include substitutions in said plurality of positions that are selected based on a function of said sequence-

30   activity relationship.

103. The method of claim 102, wherein

said modeling a sequence-activity relationship further comprises modeling a plurality of sequence-activity relationships, wherein each respective sequence-activity

relationship in said plurality of sequence-activity relationships describes the relationship between (i) one or more substitutions at one or more positions of the antibody of interest represented by the variant set and (ii) the property measured for all or said portion of the variants in the variant set; and

5   said redefining said variant set comprises redefining said variant set to comprise variants that include substitutions in said plurality of positions that are selected based on a combination function of said plurality of sequence-activity relationships.

10   104. The method of claim 103, the method further comprising:

repeating said measuring (c) using said redefined variant set, wherein a property of all or a portion of the variants in the redefined variant set is measured; and

weighting each respective sequence-activity relationship in said plurality of sequence activity relationships based on an agreement between (i) measured values

15   for the property of variants in said redefined variant set and (ii) values for the property of variants in said redefined variant set that were predicted by said respective sequence-activity relationship, wherein

a first sequence-activity relationship that achieves better agreement between measured and predicted values than a second sequence-activity relationship receives a

20   higher weight than said second sequence-activity relationship.

105. The method of claim 104 wherein said redefining further comprises redefining said variant set to comprise one or more variants each having a substitution in a position in said plurality of positions not present in any variant in the variant set

25   selected by said selecting step (b).

106. The method of claim 101 wherein said plurality of antibodies of interest represent an antibody class.

30   107. The method of claim 106 wherein said antibody class is IgG_1, IgG_2, IgG_3, IgG_4, IgA_1, or IgA_2.

108. The method of claim 101 wherein said plurality of antibodies of interest represent an antibody class and wherein said property is a level of expression of a

141

variant in a host cell, a susceptibility of a variant to a post-translational modification, a killing of a pathogenic organism or a virus resulting from an activity of said variant, a modulation of a signaling pathway by said variant, a modulation of surface density of a cell-surface receptor by said variant, a binding of a cellular growth factor receptor

5   by said variant, a binding of a receptor or a mediator of tumor-driven angiogenesis by said variant, a binding of a B cell surface antigen by said variant, a binding of a protein synthesized by said variant, an induction of an antibody-mediated cell killing by said variant, an induction of an antibody-dependent macrophage activity by said variant, an induction of a histamine release by said variant, an induction of or cross-

10   reaction with an anti-idiotype antibody by said variant, an immunogenicity of said variant, a reduction of viral titer by said variant or an immunomodulatory activity of said variant.

109. The method of claim 102, the method further comprising repeating said

15   measuring (c), modeling (d), adjusting (e), modeling (i), and redefining (ii) until a variant in said variant set exhibits a value for said property that exceeds a predetermined value.

110. The method of claim 109 wherein said predetermined value is a value that is

20   greater than the value for the property that is exhibited by said antibody of interest.

111. The method of claim 102, the method further comprising repeating said measuring (c), modeling (d), adjusting (e), modeling (i), and redefining (ii) until a variant in said variant set exhibits a value for said property that is less than a

25   predetermined value.

112. The method of claim 111 wherein said predetermined value is a value that is less than the value for the property that is exhibited by said antibody of interest.

30   113. The method of claim 101, the method further comprising repeating said measuring (c), modeling (d), adjusting (e), modeling (i), and redefining (ii) a predetermined number of times.

114. The method of claim 113 wherein said predetermined number of times is two, three, four, or five.

115. A plurality of nucleic acid sequences comprising nucleotide sequences encoding all or a portion of the variants in the redefined variant set of step (ii) of claim 109.

116. All or a portion of the variants in the redefined variant set of step (ii) of claim 109.

117. A plurality of nucleic acid sequences comprising nucleotide sequences encoding all or a portion of the variants in the redefined variant set of an instance of step (ii) of claim 111.

118. All or a portion of the variants in the redefined variant set of an instance of step (ii) of claim 111.

119. A computer program product for use in conjunction with a computer system, the computer program product comprising a computer readable storage medium and a computer program mechanism embedded therein, the computer program mechanism comprising:

a knowledge base comprising a plurality of rules; and

an expert module for constructing a variant set for an antibody of interest, the expert module comprising:

instructions for identifying, using said plurality of rules, a plurality of positions in said antibody of interest and, for each respective position in said plurality of positions, one or more substitutions for the respective position, wherein the plurality of positions and the one or more substitutions for each respective position in the plurality of positions collectively define an antibody sequence space;

instructions for selecting a variant set, wherein said variant set comprises a plurality of variants of said antibody of interest and wherein said variant set is a subset of said antibody sequence space;

instructions for measuring or receiving a measurement a property of all or a portion of the variants in said variant set;

instructions for modeling a sequence-activity relationship between (i)
one or more substitutions at one or more positions of the antibody of interest
represented by the variant set and (ii) the property measured for all or said portion of
the variants in the variant set; and

5   instructions for redefining said variant set to comprise variants that
include substitutions in said plurality of positions that are selected based on a function
of said sequence-activity relationship.

120. A computer system comprising:

10   a central processing unit;

a memory, coupled to the central processing unit, the memory storing a
knowledge base and an expert module, wherein

the knowledge base comprises a plurality of rules; and

the expert module is for constructing a variant set for an antibody of interest

15   and comprises:

instructions for identifying, using said plurality of rules, a plurality of
positions in said antibody of interest and, for each respective position in said plurality
of positions, one or more substitutions for the respective position, wherein the
plurality of positions and the one or more substitutions for each respective position in

20   the plurality of positions collectively define an antibody sequence space;

instructions for selecting a variant set, wherein said variant set
comprises a plurality of variants of said antibody of interest and wherein said variant
set is a subset of said antibody sequence space;

instructions for measuring or receiving a measurement a property of all

25   or a portion of the variants in said variant set;

instructions for modeling a sequence-activity relationship between (i)
one or more substitutions at one or more positions of the antibody of interest
represented by the variant set and (ii) the property measured for all or said portion of
the variants in the variant set; and

30   instructions for redefining said variant set to comprise variants that
include substitutions in said plurality of positions that are selected based on a function
of said sequence-activity relationship.